A Turkish Version of the Critical-Care Pain Observation Tool: Reliability and Validity Assessment

Yeşim Yaman Aktaş, PhD, RN, Neziha Karabulut, PhD, RN

Purpose: *The study aim was to evaluate the validity and reliability of the Critical-Care Pain Observation Tool in critically ill patients.*
Design: *A repeated measures design was used for the study.*
Methods: *A convenience sample of 66 patients who had undergone open-heart surgery in the cardiovascular surgery intensive care unit in Ordu, Turkey, was recruited for the study. The patients were evaluated by using the Critical-Care Pain Observation Tool at rest, during a nociceptive procedure (suctioning), and 20 minutes after the procedure while they were conscious and intubated after surgery.*
Finding: *The Turkish version of the Critical-Care Pain Observation Tool has shown statistically acceptable levels of validity and reliability. Inter-rater reliability was supported by moderate-to-high–weighted κ coefficients (weighted κ coefficient = 0.55 to 1.00). For concurrent validity, significant associations were found between the scores on the Critical-Care Pain Observation Tool and the Behavioral Pain Scale scores. Discriminant validity was also supported by higher scores during suctioning (a nociceptive procedure) versus non-nociceptive procedures. The internal consistency of the Critical-Care Pain Observation Tool was 0.72 during a nociceptive procedure and 0.71 during a non-nociceptive procedure.*
Conclusions: *The validity and reliability of the Turkish version of the Critical-Care Pain Observation Tool was determined to be acceptable for pain assessment in critical care, especially for patients who cannot communicate verbally.*

Keywords: *intensive care unit, open-heart surgery, pain, validity, reliability.*

Yeşim Yaman Aktaş, PhD, RN, Department of Surgical Nursing, Faculty of Health Sciences, Giresun University, Giresun, Turkey; and Neziha Karabulut, PhD, RN, Department of Surgical Nursing, Faculty of Health Sciences, Atatürk University, Erzurum, Turkey.

Conflict of interest: None to report.

Address correspondence to Neziha Karabulut, Department of Surgical Nursing, Faculty of Health Sciences, Atatürk University, Erzurum, Turkey; e-mail address: nezihek@mynet.com.

1089-9472/$36.00
http://dx.doi.org/10.1016/j.jopan.2015.12.015

CRITICALLY ILL PATIENTS in intensive care units (ICUs) may experience moderate to severe pain[1-3] due to acute illness, surgery, trauma, invasive or noninvasive procedures, immobility, and nursing interventions.[4] Of these nursing interventions, endotracheal suctioning, positioning, catheter placement, dressing, drain, or chest tube removal procedures, endotracheal tube removal has been previously identified to be major sources of pain.[5,6] In their study conducted with 300 critical-care patients, Eti-Aslan et al[7] reported that chest tube, endotracheal suctioning, and

dressings were the processes causing severe pain in critical-care patients.

Pain assessment is the most important step in the management of pain. It is very difficult to assess pain objectively since pain is a subjective and multidimensional concept.[8] Although pain is usually assessed with the use of self-reports,[9] critically ill patients are often unable to provide a self-report on the presence and intensity of pain, because of a reduced level of consciousness, endotracheal intubation, or the use of sedatives or muscle relaxants.[4,10] The widely used pain assessment scales such as Visual Analog Scale, Numerical Analog Scale, Faces Pain Scale, McGill Pain Questionnaire, and Brief Pain Inventory may not be appropriate assessment tools for critically ill patients due to the above-mentioned reasons in pain assessment.[11] However, other valid and reliable measures are clearly required to assess pain in nonverbal patients even though self-reporting is the most reliable way of assessing pain.[12] A number of studies have shown that behavioral assessments provide a relatively valid and reliable means of assessing pain in nonverbal patients, and various behavioral pain assessment tools have been developed, one of which is the Critical-Care Pain Observation Tool (CPOT).[4,13,14]

The CPOT was originally developed by Gélinas et al[15] based on the findings of a literature review. This tool evaluates four behavioral domains: facial expression, body movement, muscle tension, and ventilator compliance/vocalization. Items in each section are scored from 0 to 2, with a possible total score ranging from 0 to 8. Its content validity was verified with 14 critical-care nurses and physicians.[16] The CPOT was tested among different ICU groups, including cardiac surgery patients[15] and patients with a variety of diagnoses, such as trauma, surgical, and medical cases.[17]

Literature Review

The first step in providing adequate pain relief for patients is systematic and consistent assessment and documentation of pain.[13,18] Pain intensity may be quantified using behavioral-physiological scales in nonverbal patients but health care professionals' bias may influence perceptions of the patients' suffering.[19] The 2004 Thunder Project II identified behaviors displayed during procedures in 5,957 critically ill adult patients at 169 sites.[20] In this study, patients who reported pain (n = 4,278) during a procedure (ie, turning, suctioning, wound care, device removal) displayed five behaviors: grimacing (43%), rigidity (27%), wincing (24%), shutting of eyes (34%), and verbalization of complaints (24%).[20] To identify pain behaviors in critically ill patients, Gélinas et al[21] conducted a retrospective review of 183 pain episodes that occurred in the first 72 hours after the patients were intubated. Pain behaviors such as facial expressions, agitation, movement, compliance with ventilator were identified in nurses' notes 73% of the time, while physiological indicators (ie, blood pressure, heart rate, arrhythmia) were found only 24% of the time. These studies led to the development of pain measurement tools in nonverbal critically ill patients.[15,20,21]

Measurement is fundamental for nursing practice and research. A measurement is an expression of observation results by numbers after observing certain object(s) on whether they possess certain characteristics.[22] Various tools are developed and tested in the nursing discipline to evaluate health status, results of nursing interventions, or the perception of the care given. Since nursing is a scientific practice-based discipline, selection of the most appropriate measurement tool is important.[16] In this context, the purpose of the study was to provide a measurement tool in the Turkish literature to facilitate assessing pain levels of mechanically ventilated patients who experience pain frequently in ICUs. Pain is a subjective and multidimensional concept. However, development of observational or behavioral pain scoring systems is recommended since there may be no self-reporting of pain in patients in ICUs.[23]

No tool is universally accepted for use in the nonverbal patient today. Although various tools have been developed for use in nonverbal patients,[18,24] they are not used in the ICUs in Turkey.[25] This may cause inadequate pain assessment in critically ill patients.[13,14] The aim of our study was to evaluate content validity of the Turkish version of the CPOT to be used for pain assessment, assess the internal consistency reliability of the measure, and evaluate its concurrent and discriminant validity. These study results will contribute to objective pain assessments performed by critical-care nurses in Turkey.

Methods

Design

A repeated measures design was used for this study. The study used psychometric methods to test the adapted tool. To ensure the quality of the adapted scale, international norms were performed while carrying out the adaptation. The phases carried out were (1) translation, (2) content validity, and (3) psychometric testing (factor analysis, a reliability coefficient, and interitem correlations).

Participants

The data were collected between August 2012 and January 2013 after obtaining ethical approval. The sample consisted of patients who had undergone open-heart surgery in the Cardiovascular Surgery ICU, Medical Park Hospital in Ordu, Turkey. The sample size consisted of 66 patients who complied with the criteria of inclusion for the study.

The eligibility criteria were as follows: (1) age 18 years and older, (2) intubated with a need for endotracheal suctioning, (3) were conscious, as evidenced by a Ramsay Scale score of 2 or 3, and (4) the first suctioning occurred when the participant had a score of 2 and 3 on the Ramsay Scale. Exclusion criteria were patients with an ejection fraction (EF) of ≤25%, unstable hemodynamic conditions (ie, hemorrhage, delirium, etc.), received neuromuscular blockers after surgery, and received medical treatment for chronic pain. In the study, pain was evaluated before, during, and after endotracheal suctioning because this procedure was chosen as a representative major ICU nursing activity responsible for pain, as has been previously reported.[26,27]

Guadagnoli and Velicer[28] argue that when a factor has at least four loadings greater than 0.6, the analysis is reliable irrespective of sample size, although literature suggests that it is necessary to include five to 10 subjects for each scale item in studies of validity and reliability.[29,30] In the present study, many loadings were 0.60 or higher in all the factors. Thus, the sample size in this study was adequate to perform the factor analysis, although one may argue that a greater sample size is preferable. For this reason, the sample size of the research is adequate.

Instruments

The data were collected by the researcher using the "Patient Information Form," "CPOT," "RSS," and "BPS."

PATIENT INFORMATION FORM. The questionnaire form asked for demographic characteristics of the patients including age, gender, education, marital status, previous surgery, type of surgery, EF, length of stay in ICU, and duration of mechanical ventilation.

CRITICAL-CARE PAIN OBSERVATION TOOL. This BPS was developed by Gélinas et al[15] in 2006 to assess pain intensity in patients who are not able to talk and cannot express their pain verbally in ICUs. The CPOT consists of four behavioral categories: facial expression, body movement, muscle tension, and compliance with the ventilator in intubated patients or vocalization in extubated patients. The tool is a three-point Likert scale (0 to 2), with a range between 0 and 8. A lower score indicates less pain experienced by patients. The inter-rater reliability of the tool developed by Gélinas et al[15] was found as moderate to high between researcher team and critical-care nurses (weighted κ coefficient = 0.52 to 0.88). It may be associated with the sample sizes for inter-rater reliability differing for each time. In addition, the intraclass and interclass correlation coefficient (0.80 to 0.93) of the tool was found to be high.

RAMSAY SEDATION SCALE. The Ramsay Sedation Scale (RSS) was developed by Ramsay in 1974. Participants' sedation levels were measured with the use of this scale. RSS ranks levels of conscious numerically. The scale consists of six items, of which three items describe awake levels and other three items describe asleep levels as follows: (1) anxious and agitated or restless, or both; (2) cooperative, oriented, and calm; (3) responsive to commands only; (4) exhibiting brisk response to light glabellar tap or loud auditory stimulus; (5) exhibiting a sluggish response to light glabellar tap or loud auditory stimulus; and (6) unresponsive. In the scale, 1, 2, and 3 scores indicate awake levels, and 4, 5, and 6 scores indicate asleep levels. RSS is a scale from one to six, with higher levels indicating degrees of sedation.[31]

The level of sedation of patients has been assessed by RSS in the studies assessing patients' pain levels in ICUs.[15,32,33] This scale was commonly used to assess sedation in Turkey.[34,35] Therefore, RSS was preferred to determine the level of sedation of patients, and patients who had a score of 2 and 3 points were included in the study since the conscious sedation level ended at the score of 4.

BEHAVIORAL PAIN SCALE (BPS). This scale was used in our study to test the concurrent validity. BPS includes three indicators (facial expressions, movements of upper limbs, and compliance with the ventilator) that are scored from 1 to 4, with higher numbers indicating higher levels of discomfort. The total BPS score can range from 3 (no pain) to 12 (most pain).[33,34] BPS was developed by Payen et al[33] to assess pain in critically ill sedated patients in 2001. The scale was adapted to Turkish by Vatansever in 2004 and tested in critical-care patients (n = 38) who had undergone thoracotomy or abdominal surgery. The BPS scores were assessed at rest and during the nociceptive procedures (endotracheal suctioning or turning).[34] The internal consistency coefficient α was 0.80 to 0.90 in turning and 0.71 to 0.93 in endotracheal suctioning.[34]

Ethical Considerations

The study was approved by the ethics committee of the Health Sciences Institution at Ataturk University (dated: May 08, 2012, and number: 2012.2.47), and written consent was obtained from the director of the institution (dated: July 30, 2012, and number: 551). All participants were informed about the purpose of the research and were assured of their right to refuse participation or to withdraw from the study at any stage. Written consent was obtained from all participants in the study. The anonymity and confidentiality of participants was guaranteed. Written consent was obtained from Gélinas et al to adapt the CPOT.

Data Collection

The purposes and procedures of the study were explained to the patients and the medical staff. After the suitability of the patients to the criteria required for the research was evaluated, informed consent was obtained from all participants. Participants' demographic characteristics were collected at hospital admission. After the cardiac surgery, when the participant was still intubated, had become conscious, as evidenced by a Ramsay Scale score of 2 or 3 and needed suctioning in the ICU, the pain behaviors were recorded on a video camera by the researcher. Suctioning was performed by the critical-care nurse, and pain behaviors were recorded before, during and 20 minutes after suctioning by the researcher. Pain was evaluated 20 minutes after endotracheal suctioning because the stress hormones, epinephrine and norepinephrine, which both have half-lives in the 1 to 3 minutes range, are presumably released by a stressful procedure such as endotracheal suctioning but are known to return to normal levels after 15 to 20 minutes.[12] All the data collections and measurements were conducted by the first author. To test inter-rater reliability, two nurse and one physician observers performed the assessments with the CPOT independently and were blind to each other's scores. The nurses and physician who participated in reliability testing were given one educational session on the instrument, which consisted of viewing a standardized videotape of patient scenarios that was obtained from the CPOT's author. The pain assessments were performed by using a video camera. Two of the observers were bachelor degree nurses who were working in the ICU, and the third observer was a cardiovascular surgeon.

Data Analysis

The Statistical Package for the Social Sciences (SPSS, Chicago, IL) for Windows, version 18.0, was used for data entry and analysis. The patients' demographic variables were evaluated using the percentage distribution and mean. Pearson correlation was used to determine correlation scores of items and the total scale. Inter-rater reliability was examined, and weighted κ coefficients were calculated for all assessments (before, during and after suctioning). To test validity of the CPOT, we determined concurrent and discriminant validity. Concurrent validity was examined by measuring the relationship between the CPOT and BPS scores. Discriminant validity was examined at rest and during suctioning. Cronbach's alpha was calculated to find internal consistency reliability (Table 1). In the comparisons, the confidence interval was taken as 95%,

Table 1. Description of Validity and Reliability Methods Examined in This Study

Psychometric Tests	Definition*	Coefficient or Analysis	Level of Acceptability†
Criterion validity (concurrent validity)	The degree to which scores on an instrument are correlated with some external criterion In this study, Behavioral Pain Scale was used to test concurrent validity	Pearson correlation	$P < .01$
Discriminant validity	An approach used to construct validation that involves assessing the degree to which a single method measuring two distinct constructs yields different results (ie, the presence or absence of pain) In this study, we examined whether the CPOT could be used to discriminate between pain during suctioning and lack of pain at rest	Rest time compared with suctioning for all three testing periods: paired *t* test	$P < .01$
Inter-rater reliability	The degree to which two raters or observers, operating independently, assign the same ratings or values for an attribute being measured or observed Three raters assessed the patients in this study: two critical-care nurses and one physician	Weighted κ coefficients	< 0 poor 0 to 0.20 slight 0.21 to 0.40 fair 0.41 to 0.60 moderate 0.61 to 0.80 substantial 0.81 to 1.00 almost perfect
Internal consistency	The degree to which the subparts of an instrument are all measuring the same attribute or dimension, as a measure of the instrument's reliability	Cronbach's coefficient alpha	$P > .70$

*Definitions from Polit and Beck.[36]
†Levels of acceptability for inter-rater reliability scores from Landis and Koch.[37]

and a *P* value below .05 was considered to indicate a statistically significant difference.

Results

Demographic Data

A total of 66 patients were included in the study, and 198 pain behaviors of all patients in the study were observed before, during, and after suctioning. Demographic characteristics of the participants are shown in Table 2. Patients ranged from 44 to 82 years in age, with an average of 65 years. Of the study group, 72.7% were male, 86.4% were married, and 80.3% were literate or graduated from primary school. Most of the patients (89.4%) underwent coronary artery bypass grafting, and more than half (53%) stayed in the ICU for 1 day. As shown in Table 2, the mean of EF and duration of mechanic ventilation were 60.77 ± 9.62 and 4.04 ± 1.52, respectively.

Validity

In the study, the validity and reliability study of the CPOT (discriminant and concurrent validity,

Table 2. Demographic and Clinical Characteristics of the Sample (n = 66)

Variable	n	%
Gender		
Female	18	27.3
Male	48	72.7
Education		
Literate-primary school	53	80.3
High school-university	13	19.7
Marital status		
Married	57	86.4
Single	9	13.6
Previous surgery		
Yes	43	65.2
No	23	34.8
Type of surgery		
CABG	59	89.4
AVR/MVR	5	7.6
CABG (+) valve annuloplasty	2	3.0
Length of stay in ICU		
1 d	35	53.0
2 d	19	28.8
3 d	12	18.2
Age (y; mean [SD])	65.01 (1.23)	
EF (%; mean [SD])	60.77 (9.62)	
Duration of MV (h; mean [SD])	4.04 (1.52)	

EF, ejection fraction; ICU, intensive care unit; MV, mechanic ventilation; CABG, coronary artery bypass graft; AVR, aorta valve replacement; MVR, mitral valve replacement; SD, standard deviation.

inter-rater reliability) was performed in accordance with the related literature,[15,16,38] and the content validity was based on the opinions of scholars specializing in the nursing field. The following methods were used in the validity and reliability study of the CPOT. Content validity and construct validity were analyzed for the validity study, and internal consistency and inter-rater reliability were examined for the reliability study.

Translation Procedures

After obtaining the study permission, the items of the CPOT were first translated to Turkish by the researchers. The translation was also carried out by three Turkish experts, who worked independently on the translation. They were all teachers of English. The initial translation into Turkish was back translated into English by a translator, whose native language was Turkish. The translation phase had the purpose of checking for discrepancies between content and meaning of the original version and the translated instrument. All the versions were evaluated by the authors, and the final version was then formed thereby.

Content Validity

To test item clarity and content validity, the translated version was submitted to academics who were experts in the Public Health Nursing, Fundamentals of Nursing, Surgical Nursing, and the Counseling and Guidance fields. For this assessment, the Turkish form of the translated scale was given to five faculty members who were experts in their fields. Content Validity Index was used in assessing the expert opinions. The experts were informed concerning the measures and concepts involved by the authors. Based on this index, the experts were asked to evaluate each item of the scale by using a Likert-type scale[39]: 1 (not suitable), 2 (the item needs to be changed to make it suitable), 3 (suitable, but needs minor correction), and 4 (very suitable). No change was done for the rankings of the scale items. The items of the scale were based on these evaluations. The experts were in agreement for these evaluations. Conceptual adjustments were not required after translation and review.

Concurrent and Discriminant Validity

Concurrent validity and discriminant validity were investigated to test the scale validity. For concurrent validity, the "BPS" was used to evaluate pain behaviors of the patients (Table 3). According to discriminant validity, patients were assessed with the CPOT during nociceptive (suctioning) and non-nociceptive (before and after suctioning) procedures (Table 4).

As shown in Table 3, the entire correlations between CPOT scores and BPS scores were found to be statistically significant before, during, and after suctioning (weighted κ coefficient = 0.20 to 0.89). These findings show that the concurrent validity of CPOT was supported with mild to high. It is accepted levels for kappa coefficients as follows: (1) 0 to 0.20 slight; (2) 0.21 to 0.40 fair; (3) 0.41 to 0.60 moderate; (4) 0.61 to 0.80 substantial; and (5) 0.81 to 1.0 almost perfect.[16] Discriminant validity was also supported by higher scores during suctioning (a nociceptive procedure) versus non-nociceptive procedures (Table 4).

Table 3. The Correlations Between CPOT and BPS Before, During, and After Suctioning

BPS score	Before Suctioning CPOT	During Suctioning CPOT	After Suctioning CPOT
Before suctioning BPS	.836** .000		
During suctioning BPS		.891** .000	
After suctioning BPS			.851** .000

BPS, Behavioral Pain Scale; CPOT, Critical-Care Pain Observation Tool.
**$P < .001$.

Internal Consistency

In the study, item analysis was performed to determine the internal consistency of the CPOT. Internal consistency coefficient and item total correlations were calculated to test the consistency and homogeneity between CPOT items during nociceptive and non-nociceptive procedures. Cronbach's alpha coefficients were obtained for internal consistency. Internal consistency may be a necessary condition for homogeneity or unidimensionality of a scale and Cronbach's alpha should be 0.70 and more.[36] As shown in Table 5, internal consistency coefficient obtained with the CPOT scoring system was 0.72 during a nociceptive procedure and 0.71 after a nociceptive procedure. Internal consistency coefficient examines whether the items in a test or scale correlate to each other. It is accepted that the higher alpha coefficient of a scale, the more consistent items in a scale are.[36] Based on this information, it can be said that the internal consistency between the CPOT items is adequate during suctioning, which is a nociceptive procedure.

Inter-rater Reliability

Inter-rater reliability is the consistency with which two raters agree on their measurements or observations of a phenomenon.[15] The weighted kappa values are a measure of how well the ratings of the observers were in agreement. The CPOT scores were obtained by three observers independently during, before, and after a nociceptive procedure (suctioning) to test the inter-rater reliability of the CPOT.

Table 4. Differences in Scores on the Critical-Care Pain Observation Tool Before, During, and After Suctioning

Assessment Time	Mean	SD	*t*	df	*P*
Before suctioning	0.48	0.76	−12.55	65	.001
During suctioning	3.22	1.82			
During suctioning	3.22	1.82	9.372	65	.001
After suctioning	1.01	1.50			

SD, standard deviation.

Table 5. Internal Consistency of Critical-Care Pain Observation Tool

Item	Item Total Correlations
Facial expressions	
During suctioning	
r	.703
p	.000
After suctioning	
r	.823
p	.000
Body movements	
During suctioning	
r	.870
p	.000
After suctioning	
r	.837
p	.000
Muscle tension	
During suctioning	
r	.769
p	.000
After suctioning	
r	.872
p	.000
Compliance with the ventilator	
During suctioning	
r	.579
p	.000
After suctioning	
r	.607
p	.000

Table 6. Inter-rater Reliability of Critical-Care Pain Observation Tool

Observation times	First Rater	Second Rater
Facial expressions		
Before suctioning		
Second rater		
r	.868**	—
p	.000	—
Third rater		
r	.957**	.819**
p	.000	.000
During suctioning		
Second rater		
r	.775**	—
p	.000	—
Third rater		
r	.911**	.833**
p	.000	.000
After suctioning		
Second rater		
r	.874**	—
p	.000	—
Third rater		
r	.934**	.891**
p	.000	.000
Body movements		
Before suctioning		
Second rater		
r	.889**	—
p	.000	—
Third rater		
r	.945**	.940**
p	.000	.000
During suctioning		
Second rater		
r	.886**	—
p	.000	—
Third rater		
r	.845**	.933**
p	.000	.000
After suctioning		
Second rater		
r	1.000**	—
p	.000	—
Third rater		
r	1.000**	1.000**
p	.000	.000
Muscle tension		
Before suctioning		
Second rater		
r	1.000**	—
p	.000	—
Third rater		
r	.887**	.887**
p	.000	.000
During suctioning		
Second rater		
r	.806**	—
p	.000	—
Third rater		
r	.879**	.766**
p	.000	.000
After suctioning		
Second rater		
r	.891**	—
p	.000	—
Third rater		
r	.917**	.856**
p	.000	.000
Compliance with the ventilator		
Before suctioning		
Second rater		
r	1.000**	—
p	.000	—
Third rater		
r	1.000**	1.000**
p	.000	.000
During suctioning		
Second rater		
r	.546**	—
p	.000	—
Third rater		
r	.647**	.821**
p	.000	.000
After suctioning		
Second rater		
r	.702**	—
p	.000	—
Third rater		
r	.568**	.810**
p	.000	.000
Total Score		
Before suctioning		
Second rater		
r	.932**	—
p	.000	—
Third rater		
r	.935**	.884**
p	.000	.000
During suctioning		
Second rater		
r	.882**	—
p	.000	—

(Continued)

Table 6. Continued

Observation times	First Rater	Second Rater
Third rater		
r	.929**	.938**
p	.000	.000
After suctioning		
Second rater		
r	.950**	—
p	.000	—
Third rater		
r	.960**	.967**
p	.000	.000

**p< .001

As shown in Table 6, all correlation coefficients assessed by three observers were found to be statistically significant in terms of CPOT "facial expressions," "body movements," "muscle tension, "compliance with the ventilator," and "total score" before, during, and after suctioning. Inter-rater reliability of the CPOT was supported with moderate to high agreement coefficients among three observers (weighted κ coefficient = 0.55 to 1.00).

Discussion

Content validity, concurrent validity, and discriminant validity tests were used to test psychometric evaluation of the CPOT in this study. The opinions of five experts were used to assess the content validity of the scale translated to Turkish. The panel review of its content indicated that there was no need to modify its translation or content. In that case, it is likely said that content validity of the instrument has been satisfactory.

Concurrent validity and discriminant validity were investigated to test the scale validity.[16] All the correlations between the CPOT scores and BPS scores were found to be statistically significant at the third period, and patients' pain intensity scores on BPS were moderately correlated with the CPOT scores. In testing the concurrent validity of the tool developed by Gélinas et al,[38] patients' self-reports were considered, and the presence of pain was evaluated by intubated patients' nods. Gélinas et al reported that a significant association was found between patients' self-reports and the CPOT scores. Patients in pain obtained higher CPOT scores compared with patients without pain ($P \leq .001$). Our findings are consistent with the results of original tool tested in Gélinas' reference article. On the contrary, in another study conducted by Keane,[40] the findings showed a weak correlation with no significance (Spearman coefficient = 0.26; $P < .312$). This finding may be associated with small sample size (n = 23). Our results support that the criterion validity of the CPOT were mild to high because the indicators were tested against another observational pain scale.

Discriminant validity was supported by the finding that CPOT scores were significantly higher during suctioning than at rest (before and after suctioning). Similarly, Gélinas et al[15] reported that the CPOT scores were significantly higher during a nociceptive procedure (positioning; $t = -15.96$) than those at rest ($t = -9.01$). Previous studies also found higher behavioral scores during suctioning than at rest in critically ill patients.[20,26] Our study results suggest that pain behaviors are observable even if a patient cannot report pain.

Internal consistency and inter-rater reliability were assessed to test the reliability of the tool in this study. Internal consistency coefficient was 0.72 during a nociceptive procedure and 0.71 after a nociceptive procedure. Marmo and Fowler[24] tested the CPOT in patients after heart surgery and found that the tool had high reliability ($\alpha = 0.89$). The CPOT was also included in the previously study by Wibbenmeyer et al,[41] who reported a high internal consistency (Cronbach $\alpha = 0.71$) and good discriminate validity (mean scale scores = 0.27 at rest to 0.56 after noxious stimulation). On the contrary, Gélinas et al[16] reported that it was not possible to perform internal consistency and factor analysis since the tool was a one-dimensional scale of pain. In general, there is a tendency to accept 0.70 of reliability coefficient as the lower limit.[22] Based on this information, it can be said that the internal consistency of the CPOT items is high during suctioning, which is a nociceptive procedure.

Testing for inter-rater reliability in the present study showed a range of results; based on the criteria of Landis and Koch,[37] the results indicated moderate to almost perfect inter-rater reliability, with weighted kappa scores ranging from 0.55 to 1.00. Weighted kappa scores in the study by Gélinas et al[15] were similar to our findings and ranged from 0.52 to 0.88, with inter-rater reliability

ranging from moderate to high. In a repeated measures study by Vasquez et al[42] found that inter-rater reliability of the CPOT was excellent ($\kappa = 0.79$). In contrast to the present study results, Keane[40] reported that the results indicated lower inter-rater reliability scores (weighted κ coefficient = 0.34 to 1.00). Wibbenmeyer et al[41] indicated that the inter-rater reliability of the CPOT was poor (Pearson correlation coefficient, 0.63; $P < .001$). This poor reliability could be due to the limited amount of training the observers received before data collection. Scoring of a "true" facial grimace requires practice; using a standardized reference such as the Revised Faces Pain Scale may assist in helping to teach scoring of facial expression.

Study Limitations

Several factors may be considered as limitations in this study. First, the present study conducted on postoperative cardiac surgery patients, and the results should not be generalized to all critically ill patients. Second, it is suggested that a large-scale study may be conducted with other behavioral pain scoring systems such as the BPS in nonverbal patients to confirm the findings of this study.

Conclusions

This study confirmed the validity and reliability of the scale in this sample of Turkish patients who underwent open-heart surgery. The Turkish version of CPOT has shown statistically acceptable levels of validity and reliability. It is recommended that this scale should be further evaluated both in different regions of Turkey and in diverse populations. Standardized educational programs for users and clarifying the instrument in different clinical scenarios would be useful for critical-care nurses and would improve instrument reliability.

References

1. Rotondi AJ, Chelluri L, Sirio C, et al. Patients' recollections of stressful experiences while receiving prolonged mechanical ventilation in an intensive care unit. *Crit Care Med*. 2002;30:746-752.
2. Li DT, Puntillo K. A pilot study on coexisting symptoms in intensive care patients. *Appl Nurs Res*. 2006;19:216-219.
3. Cazorla C, Cravoisy A, Gibot S, Nace L, Levy B, Bollaert PE. Patients' perception of their experience in the intensive care unit. *Presse Med*. 2007;36:211-216.
4. Kwekkeboom KL, Herr KA. Assessment of pain in the critically ill. *Crit Care Nurs Clin North Am*. 2001;13:181-194.
5. Stanik-Hutt J, Soeken KL, Belcher AE, Fontaine DK, Gift AG. Pain experiences of traumatically injured patients in a critical care setting. *Am J Crit Care*. 2001;10:252-259.
6. Pasero C. Pain in the critically ill patient. *J Perianesth Nurs*. 2003;18:422-425.
7. Eti Aslan F, Badır A, Arli SK, Çakmakçı H. Patients' experience of pain after cardiac surgery. *Contemp Nurs*. 2009;34:48-54.
8. Karayurt Ö, Akyol Ö. Yoğun bakım hastalarında ağrı değerlendirmesi. [Assessment of pain in intensive care patients]. *Atatürk Üniversitesi Hemşirelik Yüksekokulu Dergisi*. 2008;11:96-104.
9. International Association for the Study of Pain (IASP). Pain terms: A list with definitions and notes on usage recommended by the IASP Subcommittee on Taxonomy. *Pain*. 1979;6:249-252.
10. Shannon K, Bucknall T. Pain assessment in critical care: What have we learnt from research. *Int Crit Care Nurs*. 2003;19:154-162.
11. Breivik H, Borchgrevink PC, Allen SM, et al. Assessment of pain. *Br J Anaesth*. 2008;101:17-24.
12. Lee K, Oh H, Suh Y, Seo W. Patterns and clinical correlates of pain among brain injury patients in critical care assessed with the critical care pain observation tool. *Pain Manag Nurs*. 2013;14:259-267.
13. Herr K, Coyne PJ, Key T, et al. Pain assessment in the nonverbal patient: Position statement with clinical practice recommendations. *Pain Manag Nurs*. 2006;7:44-52.
14. Li D, Puntillo K, Miaskowski C. A review of objective pain measures for use with critical care adult patients unable to self-report. *J Pain*. 2008;9:2-10.
15. Gélinas C, Fillion L, Puntillo KA, Viens C, Fortier M. Validation of the critical-care pain observation tool in adult patients. *Am J Crit Care*. 2006;15:420-427.
16. Gélinas C, Loiselle CG, LeMay S, Ranger M, Bouchard E, McCormack D. Theoretical, psychometric and pragmatic issues in pain measurement. *Pain Manag Nurs*. 2008;9:120-130.
17. Gèlinas C, Johnston C. Pain assessment in the critically ill ventilated adult: Validation of critical-care pain observation tool and physiologic indicators. *Clin J Pain*. 2007;23:497-505.
18. Arif-Rahu M, JoGrap M. Facial expression and pain in the critically ill non-communicative patient: State of science review. *Intensive Crit Care Nurs*. 2010;26:343-352.
19. Puntillo K. Pain assessment and management in the critically ill: Wizardry or science? *Am J Crit Care*. 2003;12:310-316.
20. Puntillo KA, Morris AB, Thompson CL, Stanik-Hutt J, White CA, Wild LR. Pain behaviors observed during six common procedures: Results from Thunder Project II. *Crit Care Med*. 2004;32:412-417.
21. Gélinas C, Fortier M, Viens C, Fillion L, Puntillo K. Pain assessment and management in critically ill intubated patients: A retrospective study. *Am J Crit Care*. 2004;13:126-135.
22. Erefe İ. *Hemşirelikte Aratırma İlke, Süreç ve Yöntemleri*. [*Nursing research principles, processes and methods*], 1. İstanbul: Odak Ofset; 2002:161-188. Baskı.

23. Herr K, Coyne PJ, McCaffery M, Manworren R, Merkel S. Pain assessment in the patient unable to self-report: position statement with clinical practice recommendations. *Pain Manag Nurs*. 2011;12:230-250.

24. Marmo L, Fowler S. Pain assessment tool in the critically ill post-open heart surgery patient population. *Pain Manag Nurs*. 2010;11:134-140.

25. Eti Aslan F, Badir A, Şelimen D. How do intensive care nurses assess patients' pain? *Nurs Crit Care*. 2003;8:62-67.

26. Arroyo-Novoa CM, Figueroa-Ramos MI, Puntillo KA, et al. Pain related to tracheal suctioning in awake acutely and critically ill adults: A descriptive study. *Int Crit Care Nurs*. 2008;24:20-27.

27. Siffleer J, Young J, Nikoletti S, Shaw T. Patients' self-report of procedural pain in the intensive care unit. *J Clin Nurs*. 2007; 16:2142-2148.

28. Guadagnoli E, Velicer WF. Relation of sample size to the stability of component patterns. *Psychol Bull*. 1988;103: 265-275.

29. Akgul A. *Tıbbi Araştırmalarda İstatistiksel Analiz Teknikleri SPSS Uygulamaları*. [*Statistical analysis techniques in medical research SPSS applications*], 2. Ankara: Emek Ofset; 2003. Baskı.

30. Davis GC, Robinson MC. A Times Mirror Compose. In: Talbot LA, ed. *Principles and Practice of Nursing Research*. St Louis, MO: Mosby; 1995.

31. Dikmen Y. *Mekanik Ventilasyon-Klinik Uygulama Temelleri*. [*Mechanical ventilation-clinical application basics*]. Ankara: Güneş Tıp Kitabevi; 2012.

32. Aïssaoui Y, Zeggwagh AA, Zekraoui A, Abidi K, Abouqal R. Validation of a behavioral pain scale in critically ill, sedated, and mechanically ventilated patients. *Anesth Analg*. 2005;101:1470-1476.

33. Payen J, Bru O, Bosson J, et al. Assessing pain in critically ill patients by using a behavioral pain scale. *Crit Care Med*. 2001;29:2258-2263.

34. Esen H, Öntürk ZK, Badır A, Eti Aslan F. Entübe ve sedatize yoğun bakım hastalarının pozisyon verme ve aspirasyon sırasındaki ağrı davranışları. [Pain behaviours of intubated and sedated intensive care patients during positioning and aspiration]. *ACU Sağlık Bil Derg*. 2010;1: 89-93.

35. Demir A, Öztürk B, Bektaş Ş, et al. Postoperative pain therapy in beating-heart coronary revascularization patients who were applied the fast-track protocol: A randomised, placebo controlled study. *Turkish J Thorac Cardiovasc Surg*. 2011;19:163-169.

36. Polit DF, Beck CT. *Essentials of Nursing Research Methods, Appraisal, and Utilization*, 6th ed. Philadelphia: Lippincott; 2005.

37. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33:159-174.

38. Gélinas C, Harel F, Fillion L, Puntillo KA, Johnston CC. Sensitivity and specificity of the critical-care pain observation tool for the detection of pain in intubated adults after cardiac surgery. *J Pain Symptom Manage*. 2009;37:58-67.

39. Aksayan S, Gözüm S. Kültürlerarası ölçek uyarlaması için rehber I: ölçek uyarlama aşamaları ve dil uyarlaması. [A guide for transcultural adaptation of the scale]. *Hemşirelik Araştırma Dergisi*. 2002;4:9-14.

40. Keane KM. Validity and reliability of the critical care pain observation tool: A replication study. *Pain Manag Nurs*. 2013; 14:e216-e225.

41. Wibbenmeyer L, Sevier A, Liao J, et al. Evaluation of the usefulness of two established pain assessment tools in a burn population. *J Burn Care Res*. 2011;32:52-60.

42. Vasquez M, Pardavila MI, Lucia M, Aguado Y, Margall MA, Asiain MC. Pain assessment in turning procedures for patients with invasive mechanical ventilation. *Nurs Crit Care*. 2011; 16:178-185.